(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 096 387 A1

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
02.05.2001 Bulletin 2001/18

(51) Int Cl.7: **G06F 13/364**

(21) Application number: **99830674.0**

(22) Date of filing: **26.10.1999**

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(71) Applicant: **BULL S.A.**
78434 Louveciennes Cedex (FR)

(72) Inventors:
- **Macchi, Alberto**
  20010 Inveruno (Milano) (IT)
- **Bosisio, Giuseppe**
  20099 Sesto San Giovanni (Milano) (IT)

(74) Representative: **Pezzoli, Ennio et al**
Jacobacci & Perani S.p.A.
Via Senato, 8
20121 Milano (IT)

(54) An arbitration unit for a bus

(57) An arbitration unit (150) for a bus comprises means (205) for receiving a request for access to the bus from each of a plurality of agents connected to the bus (130), means (240-260) for granting access to the bus to one of the requesting agents, the agent which has gained access to the bus releasing access to the bus as a result of a retry response when a resource requested is in an unavailable state, and first scheduling means (2401, 2451) for selecting one of the requesting agents by a round-robin policy, wherein there are provided means (210) for temporarily asserting a high-priority attribute for any agent which has received the retry response on a number of consecutive times greater than a threshold value, and second scheduling means (240h, 245h) for selecting one of the high-priority requesting agents by a round-robin policy, the granting means (240-260) granting access to the bus selectively to the agent selected by the first scheduling means (2401, 2451) or by the second scheduling means (240h, 245h).

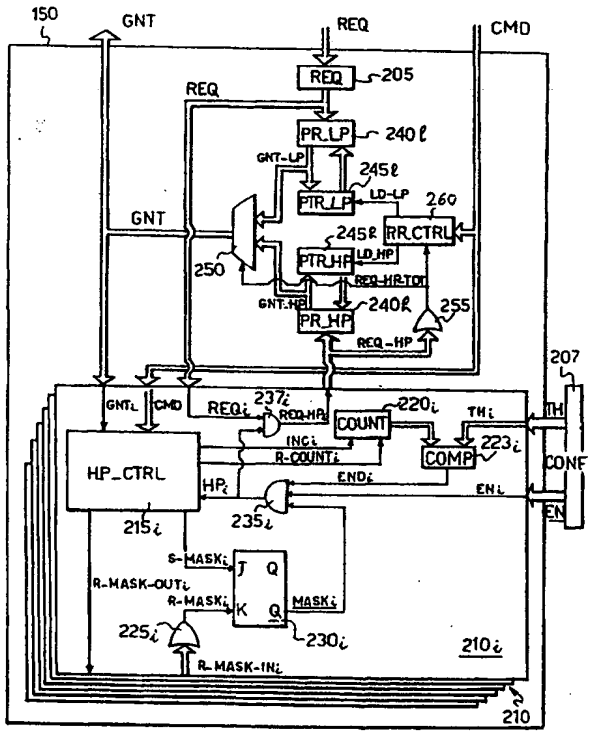

FIG.2

EP 1 096 387 A1

Description

[0001] The present invention relates to an arbitration unit for a bus.

[0002] A bus is a multi-point channel to which various units (agents) are connected in parallel for transmitting information. In general, several agents can operate in master mode, in which they control the exchange of information (transmission or receipt) with a selected agent, which operates in slave mode. Since only one agent at a time can have access to the bus in master mode, it is necessary to provide an arbitration unit which controls mutually exclusive access to the bus.

[0003] Each agent which has to access the bus sends a corresponding request to the arbitration unit. The arbitration unit typically assigns access to the bus to one of the requesting agents which is selected by a round-robin policy. In particular, the agents are arranged in a circular list of decreasing priority, starting from a position identified by a round-robin pointer which changes continuously upon each arbitration cycle; the arbitration unit thus grants access to the bus to the first requesting agent which is in a position following that identified by the round-robin pointer (in accordance with a predetermined direction of scanning of the circular list).

[0004] The bus generally provides for a retry mechanism, in which the agent which has gained access to the bus receives a retry response as a result of which it releases access to the bus (in order subsequently to repeat the access request) if the selected slave agent is temporarily unavailable to perform the required operations. This avoids the bus being kept busy unnecessarily whilst waiting for the selected slave agent to become available.

[0005] A disadvantage of this solution is that it may give rise to a situation in which an individual agent is blocked (starvation). For example, the case may be considered in which all of the agents continuously request access to the bus in order to exchange information with the same slave agent; in this situation, it is possible for one agent always to obtain access to the bus at a moment at which the slave agent is occupied by the other agents so that it continually receives the retry response.

[0006] The object of the present invention is to overcome the above-mentioned drawbacks. To achieve this object, an arbitration unit as described in the first claim is proposed.

[0007] In short, there is provided an arbitration unit for a bus comprising means for receiving a request for access to the bus from each of a plurality of agents connected to the bus, means for granting access to the bus to one of the requesting agents, the agent which has gained access to the bus releasing access to the bus as a result of a retry response when a resource requested is in an unavailable state, and first scheduling means for selecting one of the requesting agents by a round-robin policy, wherein the arbitration unit includes means for temporarily asserting a high-priority attribute for any agent which has received the retry response on a number of consecutive times greater than a threshold value, and second scheduling means for selecting one of the high-priority requesting agents by a round-robin policy, the granting means granting access to the bus selectively to the agent selected by the first scheduling means or by the second scheduling means.

[0008] Moreover, the present invention also proposes a data-processing system comprising the arbitration unit and a corresponding arbitration method.

[0009] Further characteristics and the advantages of the arbitration unit according to the present invention will become clear from the following description of a preferred embodiment thereof, given by way of nonlimiting example, with reference to the appended drawings, in which:

Figure 1 is a basic block diagram of a data-processing system in which the arbitration unit of the present invention can be used,
Figure 2 shows the structure of the arbitration unit schematically,
Figure 3 is a simplified state diagram showing the operation of a high-priority control logic of the arbitration unit.

[0010] With reference in particular to Figure 1, this shows a data-processing system 100 having an architecture with symmetrical processors. The data-processing system 100 comprises a system bus (SYSBUS) 110 to which are connected in parallel several microprocessors (μp) C1, C2, C3 and C4 (four microprocessors in the embodiment shown) and a main memory (MEM) 120; the main memory 120 is constituted by a DRAM (Dynamic Random Access Memory) with an associated respective control logic which also performs the function of arbitrating mutually exclusive access to the system bus 110.

[0011] Several peripheral input/output units (I/O) P1, P2, P3, P4 and P5 (five units in the embodiment shown) are grouped together to form a cluster. The peripheral units P1-P5 are connected to a local bus (LOCBUS) 130, for example, of the PCI (Peripheral Component Interconnect) type. A bridge unit (ITF) 140 operates as an interface between the system bus 110 and the local bus 130; in particular, the bridge unit 140 controls the different communication protocols of the system bus 110 and of the local bus 130, the different timings, and the different data and address formats. Within the bridge unit 140 there is an arbitration logic (ARB) 150 which controls mutually exclusive access to the local bus 130 by the peripheral units P1-P5 and by the bridge unit 140 (agents).

[0012] A timing unit (TEMP) 160s generates a periodic clock signal Cks which is supplied to all of the units connected to the system bus 110 (the microprocessors C1-C4, the main memory 120, and the bridge unit 140) in order to synchronise their operation. Similarly, a fur-

ther timing unit (TEMP) 1601 generates a clock signal Ckl (having a lower frequency than the clock signal Cks) which is supplied to all of the units which are connected to the local bus 130 (the peripheral units P1-P5 and the bridge unit 140).

[0013] The local bus 130 has several connection pins which are used for the exchange of signals in accordance with a corresponding communication protocol. In particular, each peripheral unit P1-P5 is connected to the arbitration logic 150 by means of a point to point dedicated line in order to transmit a signal REQi indicative of a request for access to the local bus 130 (in order to use resources of the data-processing system 100); a further access request signal REQi is supplied to the arbitration logic 150 directly within the bridge 140. The index i identifies an expansion slot in which the bridge unit 140 (i=0) or the respective peripheral unit P1-P5 (i=1...5) is housed. Each peripheral unit P1P5 is also connected to the arbitration logic 150 by means of a point to point dedicated line in order to receive a signal GNTi (where i=1...5) indicative of the granting of mutually exclusive access to the local bus 130; a further access-granting signal GNTi (where i=0) is supplied by the arbitration logic 150 directly into the bridge unit 140.

[0014] The arbitration logic 150 is connected by means of a multi-point channel, to all of the peripheral units P1-P5 and directly into the bridge unit 140 in order to receive control signals CMD of a transaction performed on the local bus 130. For example, a signal FRAME indicates the start and the duration of access to the local bus 130, a signal IRDY indicates that the agent which has obtained access to the local bus 130 (the initiator agent) is available to perform a current data phase of the transaction, a signal DEVSEL indicates that the agent selected by the initiator agent (the target agent) has recognised itself as such during an address phase of the transaction, a signal TRDY indicates that the target agent is available to perform the current data phase, a signal STOP indicates that the target agent requires the initiator agent to interrupt the transaction, etc.

[0015] The communication protocol of the local bus 130 provides for a retry mechanism in which the initiator agent receives a retry response as a result of which it releases access to the local bus 130 (in order subsequently to repeat the request for access), if the target agent is in an unavailable state and is thus temporarily unable to perform the transaction. For example, the initiator agent may receive the retry response because the target agent is blocked by another agent, has no internal resources available, etc.

[0016] For instance, the case may be considered of a writing transaction in which a peripheral unit (such as the peripheral unit P2) has gained access to the local bus 130 in order to write data in DMA (Direct Memory Access) mode onto the main memory 120 (via the bridge unit 140). The peripheral unit P2 asserts the signal FRAME (to indicate the start of the writing transaction) and then asserts the signal IRDY (to indicate that it is available to transmit the data). The bridge unit 140 asserts the signal DEVSEL (to indicate that it has recognised itself as the target of the writing transaction) and then asserts the signal STOP (to request the peripheral unit P2 to interrupt the writing transaction); the signal TRDY is still kept deasserted to indicate that the bridge unit 140 is not available to accept the data, for example, because a writing buffer is completely full. The peripheral unit P2 consequently deasserts the signal FRAME and then deasserts the signal IRDY so as to terminate the writing transaction without any transfer of data. Similarly, in the case of a reading transaction, a peripheral unit which has obtained access to the local bus 130 in order to read data in DMA mode from the main memory 120 (via the bridge unit 140) receives the retry response because the data requested is not present in a cache memory of the bridge unit 140.

[0017] These situations differ from those in which the transaction is terminated by the initiator agent when the data has been fully transferred or when access to the bus is revoked by the arbitration logic, when the transaction is aborted by the initiator agent or by the target agent because of an error, or when the transaction is terminated by the target agent owing to a disconnection request (because it is not able to supply or accept further data with the due latency).

[0018] Similar considerations apply if the data-processing system has a different architecture (even with a single processor), or has a different number of peripheral units connected to the local bus, if the local bus uses another communication protocol, if a different type of bus is considered, or if the arbitration logic is constituted by a dedicated unit, etc.

[0019] In the arbitration logic of the present invention, as described in detail below, a high-priority attribute is associated with each agent. Normally, the high-priority attribute is deasserted (the agent is in a low-priority condition). The high-priority attribute is asserted temporarily (the agent is in a high-priority condition), however, when the agent receives the retry response (without intermediate transactions terminated for other reasons) on a number of consecutive times greater than a threshold value. The arbitration logic 150 then selectively grants access to the local bus 130 to one of the low-priority requesting agents selected by a round-robin policy or to one of the high-priority requesting agents selected by a round-robin policy.

[0020] This solution ensures that the high-priority agents gain access to the local bus 130 at a greater frequency than the low-priority agents so that they have a greater probability of finding the requested resources available. The arbitration unit of the present invention thus drastically reduces the risk of starvation of the agents.

[0021] This result is achieved extremely simply; this permits the use of a very rapid structure which does not overload the arbitration unit in any way.

[0022] With reference now to Figure 2, the arbitration

logic 150 receives as inputs the access-request signals REQi (collectively indicated REQ) and the control signals CMD, and outputs the access-granting signals GNTi (collectively indicated GNT). The access-request signals REQ are latched in an input buffer 205. A bank of configuration registers (CONF) 207 of the bridge unit supplies, for each agent, a signal THi, representative of a threshold value, and an enabling signal ENi; the threshold signals THi (collectively indicated TH) and the enabling signals ENi (collectively indicated EN) are received as inputs by the arbitration logic 150.

[0023] The arbitration logic 150 includes a bank of high-priority blocks 210 (6 in the embodiment in question) each of which is associated with an agent. In particular, a generic high-priority block 210i (associated with the i-th agent) comprises a high-priority control logic (HP_CTRL) 215i, constituted basically by a state machine. The high-priority control logic 215i receives as inputs the control signals CMD and the access-granting signal GNTi.

[0024] The high-priority control logic 215i generates an increment signal INCi and a reset signal R_COUNTi which are supplied as inputs to a counter (COUNT) 220i. The value contained in the counter 220i and the threshold value THi are supplied to corresponding inputs of a comparator (COMP) 223i; the comparator 223i outputs an end-of-count signal ENDi indicative of the fact that the content of the counter 220i has reached the threshold value THi.

[0025] The high-priority logic control unit 215i also generates a self-masking deactivation signal R_MASK_OUTi which is supplied to all of the other high-priority blocks of the bank 210. The self-masking deactivation signals received by the other high-priority blocks of the bank 210 (collectively indicated R_MASK_INi) are supplied to corresponding inputs of an OR logic gate 225i. The high-priority block 210i also includes a flip-flop 230i having a set input J, a reset input K, a main output Q and a secondary output $\overline{Q}$. The input J receives a self-masking activation signal S_MASKi generated by the high-priority control logic 215i, whereas the input K receives a self-masking deactivation signal R_MASKi generated by the OR logic gate 225i.

[0026] The end-of-count signal ENDi, the enabling signal ENi and a self-masking signal $\underline{\text{MASK}}$i generated by the output $\overline{Q}$ of the flip-flop 230i are supplied to corresponding inputs of an AND logic gate 235i. The self-masking signal $\underline{\text{MASK}}$i is underlined to indicate that (unlike the other signals) it is at a logic level 0 when asserted and at a logic level 1 when deasserted. The AND logic gate 235i generates a signal HPi indicative of a high-priority attribute of the i-th agent which is supplied as an input to the high-priority control logic 215i. The high-priority signal HPi and the access request signal REQi are supplied to corresponding inputs of a further AND logic gate 237i. The AND logic gate 237i generates a high-priority access request signal REQ_HPi which is supplied outside the high-priority block 210i.

[0027] The access-request signals REQ (latched in the input buffer 205) are also supplied as inputs to a low-priority round-robin logic (RR_LP) 240l, constituted, for example, by a plurality of fixed-priority combinatorial networks connected to a multiplexer; associated with the round-robin logic 240l is a register 245l in which a corresponding low-priority round-robin pointer PTR_LP is stored. The round-robin pointer PTR_LP is supplied as an input to the round-robin logic 240l which generates a low-priority access-granting signal GNT_LPi for each agent; the access-granting signals GNT_LPi (collectively indicated GNT_LP) are supplied as inputs to the register 245l.

[0028] Similarly, the access-request signals REQ_HPi generated by each high-priority block of the bank 210 (collectively indicated REQ_HP) are supplied as inputs to a high-priority round-robin logic (RR_HP) 240h, with which is associated a register 245h wherein a corresponding high-priority round-robin pointer PTR_HP is stored. The round-robin pointer PTR_HP is supplied as an input to the round-robin logic 240h which generates a high-priority access-granting signal GNT_HPi for each agent; the granting signals GNT_HPi (collectively indicated GNT_HP) are supplied as inputs to the register 245h.

[0029] The access-granting signals GNT_LP and the access-granting signals GNT_HP are supplied to corresponding data inputs of a multiplexer 250 which outputs the access-granting signals GNT. The access-request signals REQ HP are also supplied to corresponding inputs of an OR logic gate 255; the OR logic gate 255 generates a total high-priority access request signal REQ_HP_TOT which is supplied to a control input of the multiplexer 250.

[0030] A round-robin control logic (RR_CTRL) 260 receives as inputs the control signals CMD and the total high-priority access-request signal REQ_HP_TOT. The round-robin control logic 260 generates two loading signals LD_LP and LD_HP which are supplied to the register 245l and to the register 245h, respectively.

[0031] The round-robin logic 240l selects one of the agents which has requested access to the local bus (access-request signal REQi asserted). In particular, the agents are arranged in a circular list defined by the corresponding indices with the last agent (i=5) followed by the first agent (i=0); the round-robin logic 240l asserts the access-granting signal GNT_LPi associated with the first requesting agent which is in a position following that identified by the round-robin pointer PTR_LP, in accordance with an increasing wrap-around sequence.

[0032] Similarly, the round-robin logic 240h selects one of the high-priority requesting agents (access request signal REQ_HPi asserted); in particular, the round-robin logic 240h asserts the access-granting signal GNT_HPi associated with the first high-priority requesting agent which is in a position following that identified by the round-robin pointer PTR_HP, in accordance with a decreasing wrap-around sequence.

[0033] At the same time, if at least one of the access-request signals REQ_HP is asserted, the OR logic gate 255 asserts the total high-priority access request signal REQ_HP_TOT. For each arbitration cycle, if the total high-priority access-request signal REQ_HP_TOT is asserted, the multiplexer 250 transfers the access-granting signals GNT_HP to its own output (so as to grant access to the local bus to the requesting agent selected by the round-robin logic 240h); at the same time, the round-robin control logic 260 asserts the loading signal LD_HP, so as to update the round-robin pointer PTR_HP to the value which identifies the agent selected. In the opposite case, the multiplexer 250 transfers the access-granting signals GNT_LP to its own output (so as to grant access to the local bus to the agent selected by the round-robin logic 240l); at the same time, the round-robin control logic 260 asserts the loading signal LD_LP so as to update the round-robin pointer PTR_LP.

[0034] In the above-described structure, each high-priority requesting agent can be selected either by the round-robin logic 240l or by the round-robin logic 240h; moreover, access to the local bus is granted to the requesting agent selected by the round-robin logic 240l only when there is no high-priority requesting agent.

[0035] This solution ensures that every high-priority requesting agent gains access to the local bus at an extremely high frequency. In particular, if the writing buffer of the bridge unit has a fairly large depth, it is very probable that at most only one high-priority requesting agent will be present; in this situation, the single high-priority requesting agent can gain access to the local bus even several times in succession, since it can be selected either by the round-robin logic 240h (the single high-priority requesting agent) or by the round-robin logic 240l.

[0036] Moreover, the fact that the circular lists controlled by the round-robin logic 240l and by the round-robin logic 240h are scanned on the basis of two opposite directions of scanning (increasing and decreasing, respectively) prevents beating phenomena, further reducing the possibility of starvation of the agents.

[0037] Similar considerations apply if the round-robin logic (low-priority and high-priority) is implemented differently and if the round-robin pointer is updated by a criterion other than the skipping criterion described above, for example, by a pseudo-random or preset-increment criterion (a conventional round-robin) or, more generally, if other equivalent scheduling means are provided, if the circular lists are scanned in a different manner, etc. Alternatively, each high-priority requesting agent may be selected solely by the round-robin logic 240h (with the corresponding access-request signal REQi masked for the round-robin logic 240l), access to the local bus may be granted by alternating the agents selected by the round-robin logic 240l and by the round-robin logic 240h in a different manner (for example, by granting access to the local bus on two consecutive occasions to the agent selected by the round-robin logic 240h), or the circular lists may be scanned in the same direction.

[0038] In order to describe the operation of the high-priority block 210i, Figure 2 is considered in combination with Figure 3 which shows a simplified state diagram 300 of the high-priority control logic 215i which progresses between various stable states upon each clock signal Ckl (not shown in the drawings for simplicity of illustration).

[0039] In an initial condition (in response to a reset signal), the high-priority control logic 215i is brought to an inactive state (IDL) 305. At the same time, the flip-flop 230i is brought to a reset state so that the self-masking signal MASKi generated by the output Q is deasserted (logic level 1). The counter 220i is reset; the content of the counter 220i is less than the threshold value THi so that the end-of-count signal ENDi generated by the comparator 223i is deasserted. The signal HPi is thus deasserted and the i-th agent is therefore in the low-priority condition. The round-robin pointer PTR_LP (stored in the register 2451) and the round-robin pointer PTR_HP (stored in the register 245h) are set at a default value (for example 0). Upon each clock signal Ckl in which the i-th agent does not gain access to the local bus, indicated by the condition NOT GNTi, the high-priority control logic 215i remains in the inactive state 305.

[0040] As soon as the i-th agent (in the low-priority condition) gains access to the local bus, indicated by the condition GNTi AND (NOT HPi), the high-priority control logic 215i changes to the state 310 (LP) and simultaneously asserts the self-masking deactivation signal R_MASK_OUTi (which is supplied to the other high-priority blocks of the bank 210).

[0041] If the i-th agent receives the retry response, indicated generically by the condition RETRY, the high-priority control logic 215i returns to the inactive state 305 and at the same time asserts the increment signal INCi. If, on the other hand, the i-th agent does not receive the retry response, indicated by the condition NOT RETRY, the high-priority control logic 215i returns to the inactive state 305 and simultaneously asserts the reset signal R_COUNTi. The counter 220i is thus incremented upon each retry response and is reset upon each non-retry response, so that the content of the counter 220i represents the number of consecutive retry responses for the i-th agent.

[0042] This situation is maintained as long as the content of the counter 220i is below the threshold value THi. As soon as the content of the counter 220i reaches the threshold value THi, the end-of-count signal ENDi generated by the comparator 223i is asserted. In this situation (upon the assumption that the enabling signal ENi is asserted) the signal HPi is asserted so that the i-th agent is brought to the high-priority condition. As soon as the i-th agent gains access to the local bus again, indicated by the condition GNTi AND HPi, the high-priority control logic 215i changes to the state 315 (HP).

[0043] If the i-th agent receives the retry response

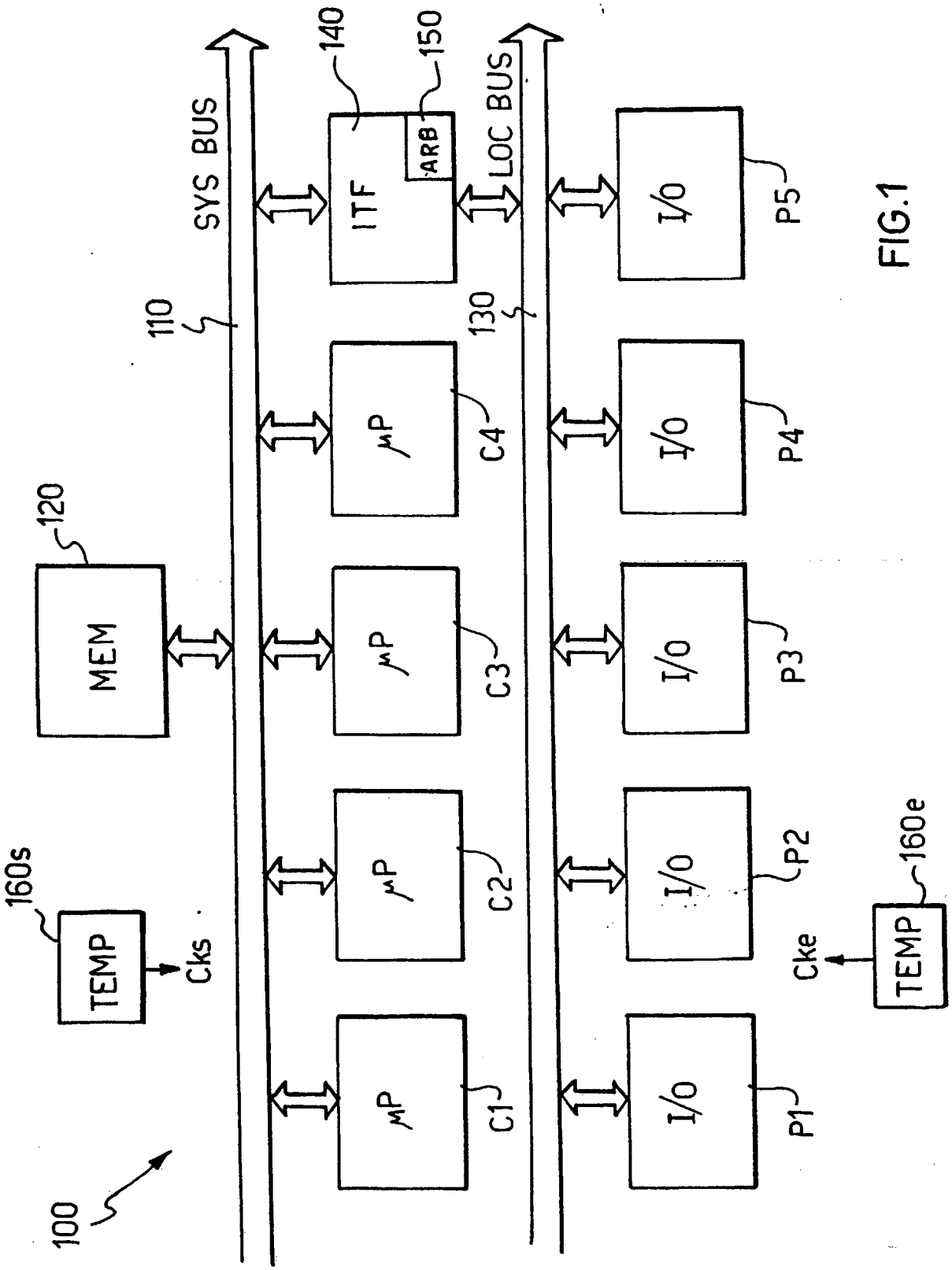
100
SYS BUS
LOC BUS
110
130
140
150
ITF
ARB
MEM
120
μP
C1
C2
C3
C4
I/O
P1
P2
P3
P4
P5
TEMP
160s
Cks
TEMP
160e
Cke

FIG.1

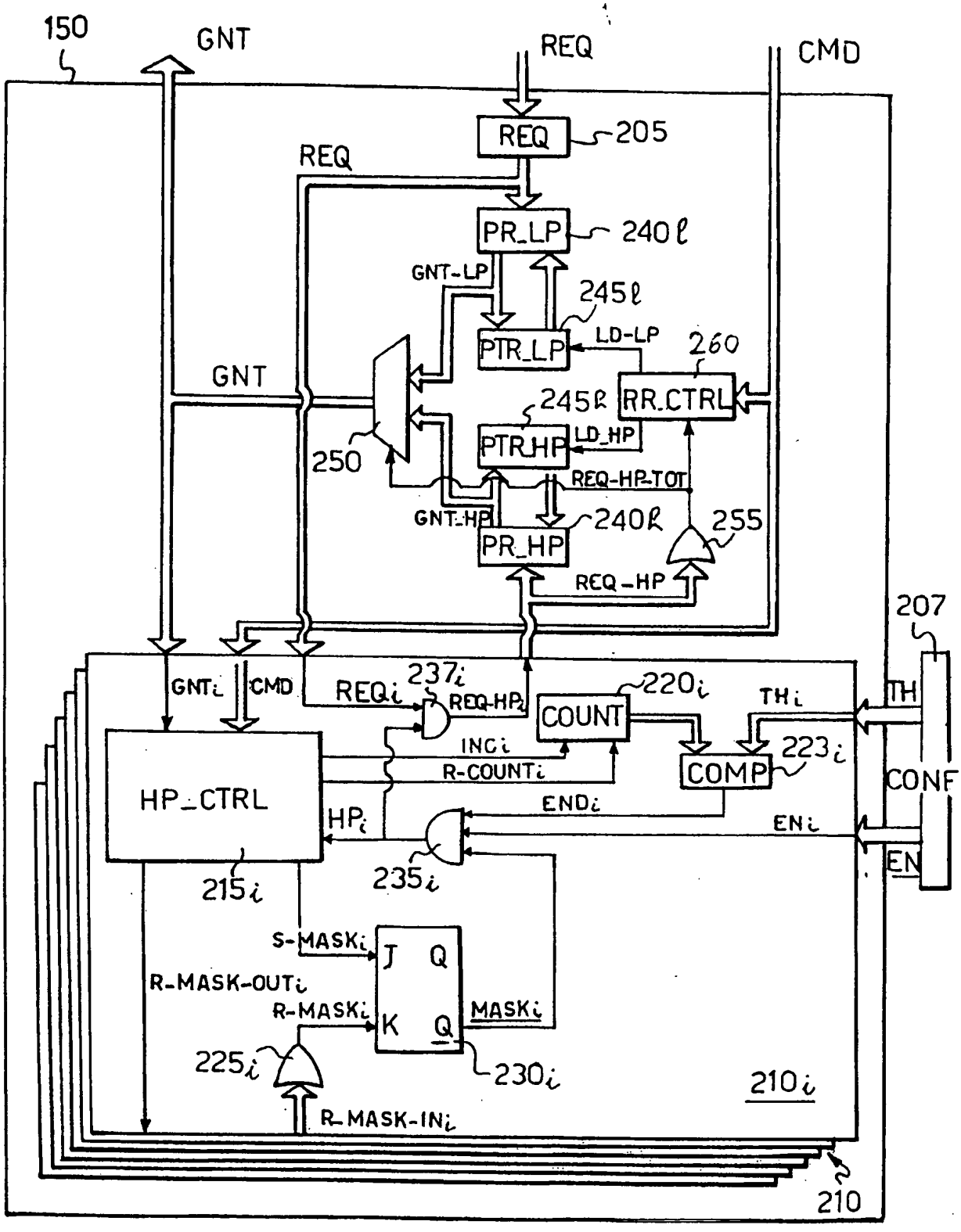
150
GNT
REQ
CMD
REQ
205
REQ
PR_LP
240ℓ
GNT-LP
245ℓ
PTR_LP
LD-LP
260
GNT
RR_CTRL
245ℛ
PTR_HP
LD_HP
250
REQ-HP-TOT
GNT-HP
240ℛ
PR_HP
255
REQ-HP
207
GNTi
CMD
REQi
237i
REQ-HPi
COUNT
220i
THi
TH
INCi
R-COUNTi
COMP
223i
CONF
HP_CTRL
HPi
ENDi
ENi
EN
215i
235i
S-MASKi
J Q
R-MASK-OUTi
R-MASKi
K Q
MASKi
225i
230i
210i
R-MASK-INi
210

FIG.2

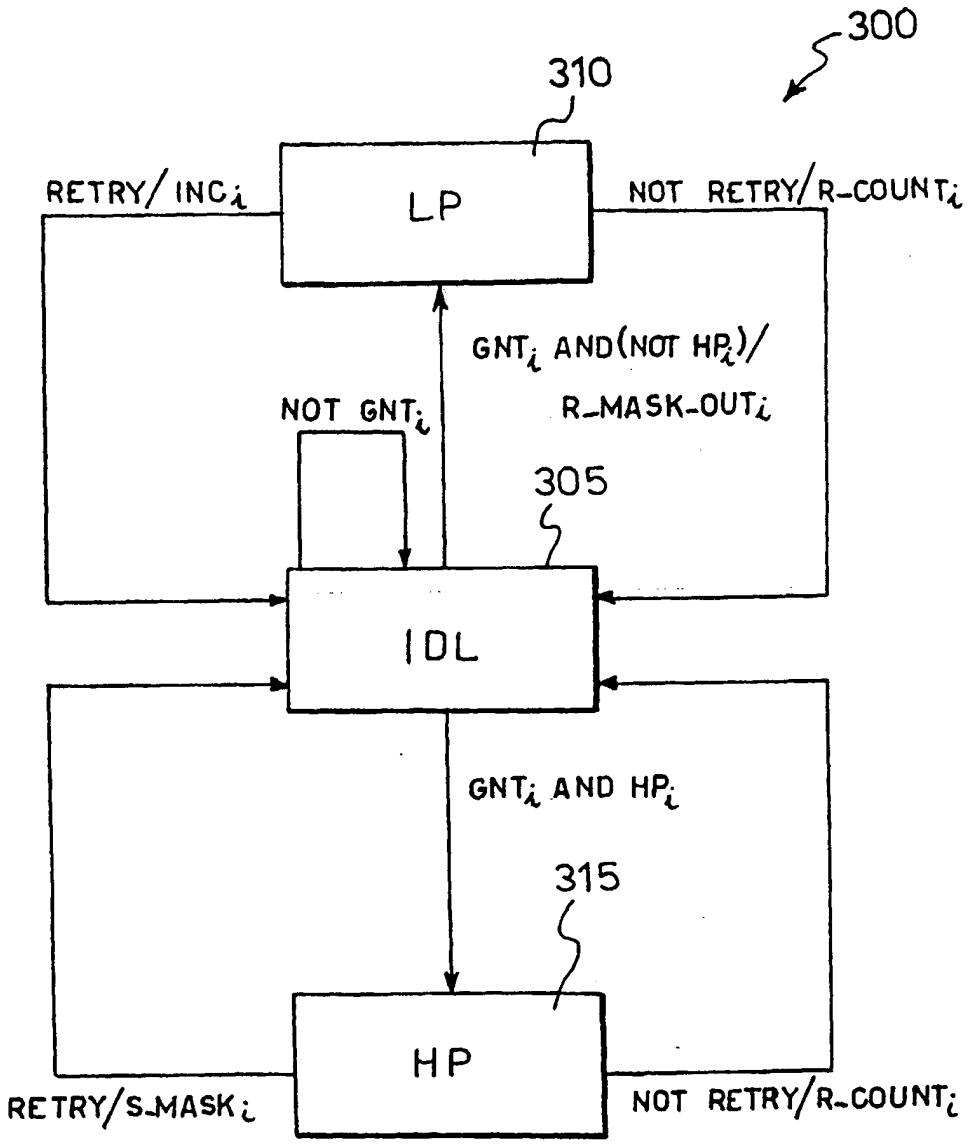
300
310
LP
RETRY/INC$_i$
NOT RETRY/R-COUNT$_i$
GNT$_i$ AND(NOT HP$_i$)/
R_MASK-OUT$_i$
NOT GNT$_i$
305
IDL
GNT$_i$ AND HP$_i$
315
HP
RETRY/S_MASK$_i$
NOT RETRY/R-COUNT$_i$

FIG.3

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 99 83 0674

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
X	US 5 644 733 A (KALISH ET AL) 1 July 1997 (1997-07-01) * column 9, line 1 - column 10, line 43 * * claims 1-3; figures 1,4 *	1-3,6,7, 9,10	G06F13/364
A	EP 0 665 501 A (COMPAQ COMPUTER CORPORATION) 2 August 1995 (1995-08-02) * page 2, line 56 - page 3, line 9 * * page 5, line 2 - page 7, line 42 * * claims 1-5; figures 4,5 *	1-10	
A	US 5 889 972 A (ALLINGHAM) 30 March 1999 (1999-03-30) * column 3, line 1 - column 4, line 13 * * claims 1-10; figures 2,4 *	4-7	
			TECHNICAL FIELDS SEARCHED (Int.Cl.7) G06F

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	12 April 2000	McDonagh, F

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 99 83 0674

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

12-04-2000

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5644733 A	01-07-1997	NONE	
EP 665501 A	02-08-1995	CA 2140685 A	29-07-1995
		JP 2596894 B	02-04-1997
		JP 7219893 A	18-08-1995
		US 5797020 A	18-08-1998
US 5889972 A	30-03-1999	NONE	

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82